

Soumission

Journal de la Société Française de Statistique

Analyse de sensibilité globale d'un modèle spatialisé pour l'évaluation économique du risque d'inondation

Title: Spatial global sensitivity analysis: case study on the economic modelling of flood risk

Nathalie Saint-Geours^{1,2}, Christian Lavergne², Jean-Stéphane Bailly¹ et Frédéric Grelot³

Résumé : L'analyse de sensibilité globale peine à se développer dans le champ de la modélisation environnementale. Dans sa formulation initiale, elle est limitée à l'étude de modèles $Y = f(X_1, \dots, X_p)$ où les variables d'entrée X_j et la sortie Y sont scalaires, alors que nombre de modèles environnementaux incluent une dimension spatiale marquée, soit qu'ils fassent appel à des cartes comme variables d'entrée, soit que leurs sorties soient distribuées spatialement. Au travers d'une étude de cas détaillée, nous présentons dans cet article une extension de l'analyse de sensibilité globale à l'étude de modèles spatialisés. Le modèle étudié, nommé ACB-DE, est un outil d'évaluation économique du risque d'inondation. Il est ici appliqué sur la basse-vallée de l'Orb (Hérault). Des spécifications spatialisées de l'incertitude sont utilisées pour générer un nombre fini de réalisations aléatoires équiprobables des variables d'entrée qui sont des cartes : les effets de structure spatiale ou d'auto-corrélation dans ces cartes peuvent ainsi être pris en compte. La réalisation de cartes d'indices de sensibilité permet ensuite d'étudier les sorties spatialisées du modèle ACB-DE et de rendre compte de la variabilité spatiale des indices de Sobol'. L'influence relative des variables d'entrée à différentes échelles d'étude est analysée par la réalisation de cartes d'indices de sensibilité de résolution croissante. L'analyse réalisée permet d'identifier les variables d'entrée incertaines qui expliquent la plus grande part de la variabilité de l'indicateur économique fourni par le modèle ACB-DE ; elle apporte un éclairage nouveau sur le choix de l'échelle adéquate de représentation spatialisée de cet indicateur selon la précision des variables d'entrée. L'approche proposée pourrait être aisément appliquée à d'autres modèles spatialisés peu coûteux en temps de calcul.

Abstract: Variance-based Sobol' global sensitivity analysis (GSA) was initially designed for the study of models with scalar inputs and outputs, while many models in the environmental field are spatially explicit. As a result, GSA is not a common practise in environmental modelling. In this paper we describe a detailed case study where GSA is performed on a spatially dependent model for flood risk economic assessment on the Orb valley (southeast France). Spatial input factors are handled by associating randomly generated map realizations to scalar values sampled from discrete uniform distributions. The realisations of random input maps can be generated by any method including geostatistical simulation techniques, allowing for spatial structure and auto-correlation to be taken into account. The estimation of sensitivity indices on ACB-DE spatial outputs makes it possible to produce maps of sensitivity indices. These maps describe the spatial variability of Sobol' indices. Sensitivity maps of different resolutions are then compared to discuss the relative influence of uncertain input factors at different scales.

Mots-clés : sensibilité, Sobol, spatial, inondation

Keywords: sensitivity, variance-based, Sobol, spatial, flood

Classification AMS 2000 : primaire 62P12

¹ AgroParisTech, UMR TETIS.

E-mail : saintge@teledetection.fr and E-mail : bailly@teledetection.fr

² Institut de Mathématiques et de Modélisation de Montpellier.

E-mail : christian.lavergne@univ-montp3.fr

³ Cemagref, UMR G-EAU.

E-mail : frederic.grelot@cemagref.fr

1. Introduction

Nous présentons dans cet article une application de l'analyse de sensibilité globale basée sur la variance à un modèle dont les variables d'entrée et les sorties sont distribuées spatialement. L'estimation des indices de Sobol est une approche communément développée pour mener l'analyse de sensibilité de modèles de type *boîte noire* peu coûteux en temps de calcul ; elle s'appuie sur la décomposition de la variance de Y (sortie du modèle) en variances conditionnelles. Dans sa formulation initiale, elle est limitée à l'étude de modèles $Y = f(X_1, \dots, X_p)$ où les variables d'entrée X_j sont scalaires, tout comme la sortie Y . Or, dans le vaste champ de la recherche en environnement, notamment dans l'étude des risques naturels, nombre de modèles incluent une dimension spatiale marquée, soit qu'ils fassent appel à des *cartes* comme variables d'entrée, soit que leurs sorties soient spatialisées (cartes de risque par exemple). De ce fait, l'analyse de sensibilité basée sur la variance peine à se développer dans ces champs thématiques [4]. Pourtant, une prise en compte explicite des incertitudes dans la modélisation et l'évaluation des risques naturels est indispensable : des décisions importantes concernant l'aménagement du territoire peuvent en effet être prises à l'aide de ces outils. Est-il donc possible d'adapter l'analyse de sensibilité globale à l'étude de modèles spatialisés ? Quel peut être le potentiel d'une telle approche dans la compréhension des aspects proprement *spatiaux* de ces modèles ?

Plusieurs publications apportent des éléments de réponse, que l'on discute dans un bref état de l'art en section 2. On se propose ensuite de combiner deux des approches présentées pour mener à bien l'analyse de sensibilité d'un modèle spatialisé d'évaluation économique du risque d'inondation (modèle ACB-DE, décrit en section 3). On veut répondre à des questions telles que « *Quelles variables d'entrée incertaines expliquent la plus grande part de la variabilité des sorties du modèle ACB-DE ?* », « *Quelle est la variabilité spatiale de la sensibilité des variables d'entrée ?* » et « *Comment l'échelle d'étude influence-t-elle les résultats de l'analyse de sensibilité ?* ». L'approche proposée semble originale dans la mesure où elle combine le traitement de variables d'entrée spatialisées (selon la méthode de Lilburne et Tarantola [13]) et de sorties spatialisées (via des cartes de sensibilité comme proposé par Marrel et al [14]) dans le cadre de l'analyse de sensibilité globale. Elle explore de plus un point non abordé à notre connaissance, qui est l'impact de l'échelle d'étude sur la valeur des indices de Sobol. Elle se limite au cadre de modèles peu coûteux en temps de calcul et ne fait donc pas appel à de la méta-modélisation. Sa mise en œuvre sur le modèle ACB-DE est présentée en détails en section 4.

Les résultats (section 5) illustrent la faisabilité et l'intérêt de l'approche proposée, qui offre un cadre simple pour l'analyse de sensibilité globale de modèles spatialisés. Il est ainsi possible :

- d'intégrer des variables d'entrées spatialisées à l'analyse de sensibilité et de spécifier l'incertitude pesant sur chacune d'elles de manière sophistiquée (prise en compte de la corrélation spatiale par exemple) ;
- d'établir des cartes de sensibilité pour une sortie Y spatialisée, afin de discuter de la variabilité spatiale des indices de Sobol ;
- de comparer l'influence relative des variables d'entrée à différentes échelles, par la réalisation de cartes de sensibilité de résolution croissante.

Ces différents points, ainsi que les limites de ce travail et les suites à lui donner, sont discutés en section 6.

2. Analyse de sensibilité globale spatialisée : état de l'art

Plusieurs travaux ont déjà apporté des éléments de réponse aux problèmes posés par l'analyse de sensibilité globale d'un modèle spatialisé ; ils se sont intéressés à l'intégration de variables d'entrée spatialisées et à l'étude de sorties spatialisées. On en fait ici un rapide tour d'horizon (2.1 et 2.2), puis l'on présente et l'on justifie les choix qui ont été faits pour notre étude de cas (2.3).

2.1. Méthodes pour l'intégration d'une variable d'entrée spatialisée

L'analyse de sensibilité globale telle que présentée dans Sobol et al. [21] suppose que les variables d'entrée X_i sont des variables aléatoires réelles (et indépendantes). Plusieurs travaux ont cherché à lever cette limitation pour intégrer une variable d'entrée Z distribuée spatialement.

Groupe de scalaires corrélés Une première approche peut être de considérer une variable d'entrée spatialisée Z comme un ensemble de variables scalaires $Z^{(j)}$ (par exemple les valeurs d'un champ spatial en tout pixel d'une grille régulière). Cependant, s'il existe une structure spatiale dans la variable d'entrée Z , alors les variables scalaires $Z^{(j)}$ qui la composent seront corrélées : elles ne pourront être intégrées à l'analyse de sensibilité que sous la forme d'un « groupe de variables », et un indice de Sobol S_Z sera estimé pour l'ensemble du groupe (voir le paragraphe 1.2.15 de Saltelli et al. [19] pour une discussion sur l'analyse de sensibilité globale par groupes). Cette approche, utilisée par exemple par Heuvelink et al. [9], devient impraticable dès lors que le nombre K de variables scalaires constituant la variable d'entrée spatialisée Z est trop grand ($K \gg 1000$).

Réduction de la dimension Une approche similaire est de résumer l'information contenue dans la variable d'entrée spatialisée Z en un nombre restreint de variables scalaires $Z^{(1)}$ à $Z^{(k)}$ que l'on peut considérer indépendantes. Ce sont alors ces variables scalaires $Z^{(j)}$ qui sont échantillonnées indépendamment et pour lesquelles on estime des indices de sensibilité distincts. Cette méthode est employée par Volkova et al. [22] (décomposition de l'aire d'étude en quelques zones indépendantes décrites par des paramètres scalaires) et par Lamboni (section 4.2.4 de [11]) (décomposition de Karhunen-Loève de la variable d'entrée spatialisée Z et estimation des indices de sensibilité des variables aléatoires ξ_i centrées, réduites et non corrélées issues de cette décomposition).

D'autres approches choisissent d'utiliser une spécification plus élaborée de l'incertitude qui pèse sur la variable d'entrée spatialisée Z , en mobilisant par exemple des méthodes géostatistiques, afin de tenir compte de la structure spatiale de l'incertitude.

Variable interrupteur L'approche retenue par Crosetto et al. [3] est la suivante : on introduit dans l'analyse de sensibilité une « variable interrupteur » (trigger parameter), notée E , qui suit une loi uniforme dans $[0; 1]$. Lorsque $E < 0.5$, la variable d'entrée spatialisée Z conserve sa valeur nominale et lorsque $E > 0.5$, un champ d'erreur aléatoire est ajouté à la variable d'entrée spatialisée. Ce champ d'erreur peut avoir une structure de covariance spatiale. L'indice de sensibilité S_E de la variable scalaire E est alors censé refléter l'influence de l'incertitude de la variable d'entrée spatialisée Z sur la variabilité de la sortie Y . Des réserves ont cependant été émises sur cette méthode, car l'introduction de la variable intermédiaire E rend l'interprétation de l'indice de Sobol S_E difficile.

Carte scénario Ruffo et al. [16] proposent de générer en amont de l'analyse de sensibilité un petit nombre de réalisations aléatoires de la variable d'entrée spatialisée Z par simulation géostatistique. Ces réalisations sont alors considérées comme des *scénarii* et une analyse de sensibilité distincte, prenant en compte les variables d'entrée autres que Z , est menée pour chaque scénario.

Lilburne & Tarantola Lilburne & Tarantola [13] proposent de générer, en amont de l'analyse de sensibilité, un jeu de n réalisations aléatoires de la variable d'entrée spatialisée Z (n possiblement grand). Ces n réalisations sont considérées équiprobables, et chacune est associée à un unique entier entre 1 et n . On introduit alors dans l'analyse de sensibilité une variable d'entrée scalaire Z' prenant ses valeurs de manière équiprobable dans l'ensemble $\llbracket 1, n \rrbracket$. La valeur prise par Z' indique le *numéro de la réalisation aléatoire* de la variable d'entrée spatialisée Z à utiliser pour simuler le modèle. L'indice de sensibilité $S_{Z'}$ de la variable d'entrée Z' traduit alors l'influence de l'incertitude de la variable d'entrée spatialisée Z sur la variabilité de la sortie Y .

Méta-modèles joints Enfin, Iooss & Ribatet [10] proposent de considérer une variable d'entrée spatialisée Z comme une variable d'entrée fonctionnelle « *incontrôlable* », opposée aux variables scalaires $\mathbf{X} = (X_1, \dots, X_p)$. Un méta-modèle sur l'espérance conditionnelle $E[Y(\mathbf{X}, Z) | \mathbf{X}]$ permet d'estimer les indices de sensibilité de premier ordre S_i des variables d'entrée scalaires X_i , tandis qu'un méta-modèle sur la dispersion conditionnelle $\text{Var}[Y(\mathbf{X}, Z) | \mathbf{X}]$ permet d'estimer l'indice de sensibilité total ST_Z de la variable d'entrée fonctionnelle Z . Le cas où plusieurs variables d'entrée fonctionnelles Z_j sont présentes peut être traité en prenant Z comme l'ensemble de ces variables, mais l'on ne peut alors pas obtenir un indice de sensibilité distinct ST_j pour chaque Z_j .

2.2. Méthodes pour l'analyse d'une sortie spatialisée

La prise en compte d'une sortie Y spatialisée dans une analyse de sensibilité globale a été peu explorée. Marrel et al. [14] se placent dans le cas où la sortie spatialisée Y du modèle est une grille régulière donnant en tout pixel u une valeur réelle $Y(u)$. Ils proposent d'utiliser des *cartes d'indices de sensibilité*, où est représenté, en tout point u de la grille et pour chaque variable d'entrée X_i , l'indice de sensibilité « *ponctuel* » S_i^u qui décrit l'influence de X_i sur la sortie $Y(u)$. D'autres travaux se sont intéressés à la question plus générale d'une sortie Y multivariée : Lamboni et al. [12] définissent dans ce cadre des « *indices de sensibilité généralisés* », qui n'ont pour l'instant été appliqués qu'au cas d'une sortie dynamique $Y(t)$.

2.3. Méthode retenue pour l'étude de l'outil ACB-DE

On propose dans la suite de ce papier de combiner deux des approches présentées pour mener à bien l'analyse de sensibilité du modèle ACB-DE : on associe le traitement des variables d'entrée spatialisées selon la méthode de Lilburne et Tarantola [13] et l'analyse de sorties spatialisées via des cartes de sensibilité comme proposé par Marrel et al [14]. La méthode proposée explore de plus un point non abordé à notre connaissance, qui est l'impact de l'échelle d'étude sur la valeur des indices de Sobol : on construit pour ce faire des cartes d'indices de sensibilité à des résolutions spatiales croissantes. La mise en œuvre de cette approche sur le modèle ACB-DE est présentée en détails en section 4.

Le choix de la méthode de Lilburne & Tarantola pour intégrer les variables d'entrée spatialisées s'appuie sur plusieurs arguments. Premièrement, dans notre étude de l'outil ACB-DE sur la basse-vallée de l'Orb, l'une des variables d'entrée spatialisées est représentée par une grille de 3500×3500 pixels, soit un total de 12250000 pixels : la dimension de cette variable rend impraticable une approche de type « *groupe de scalaires corrélés* ». Deuxièmement, la méthode « *réduction de la dimension* » est ici peu adaptée étant donnée la nature multiple de l'incertitude pesant sur l'une des variables d'entrée spatialisée (carte d'occupation du sol), qui rend difficile sa réduction à quelques paramètres scalaires. Troisièmement, le modèle ACB-DE présente trois variables d'entrée distribuées spatialement dont on veut mesurer la sensibilité séparément : on ne peut donc pas utiliser ici l'approche « *méta-modèles joints* » proposée par Iooss & Ribatet [10]. Enfin, par rapport à l'approche par scénarii proposée par Ruffo et al. [16], la méthode de Lilburne & Tarantola présente l'avantage de pouvoir prendre en compte un grand nombre de réalisations aléatoires de la variable d'entrée spatialisée Z et donc de mieux couvrir l'espace des réalisations possibles (alors que le nombre de scénarii est limité chez Ruffo et al.), et permet de quantifier l'influence de la variable d'entrée spatialisée Z et de ses interactions avec les autres variables.

3. ACB-DE : un outil spatialisé pour l'évaluation économique du risque d'inondation

Cette section présente brièvement le modèle ACB-DE et précise les variables d'entrée et les sorties considérées dans l'analyse de sensibilité.

3.1. Présentation

L'outil ACB-DE (Analyse Coût-Bénéfice - Dommages Évités) est un outil spatialisé d'évaluation économique des politiques de prévention des inondations. Il s'appuie sur une démarche générique d'analyse coût-bénéfice des mesures de gestion d'un risque naturel, où les bénéfices attendus d'une mesure sont estimés par les dommages qu'elle permet d'éviter. Cette méthode est générale et son implémentation peut différer d'une application à l'autre : nous utilisons dans ce papier les spécifications de l'outil ACB-DE présentées par Erdlenbruch et al [7].

L'outil ACB-DE fournit, entre autres sorties, un indicateur des *Dommages Évités Moyens Annualisés*, noté DEMA, qui permet de comparer l'efficacité économique de différents aménagements de réduction du risque d'inondation (mise en place de digues, de barrages, etc. . .). Cet indicateur peut être spatialisé, afin de localiser sur le territoire étudié les bénéfices apportés par un aménagement. Le calcul de l'indicateur DEMA s'appuie sur l'estimation des dommages potentiels dus à différents scénarios d'inondation (par exemple une crue décennale, trentennale, cinquantennale, centennale et une crue extrême). Pour estimer les dommages dus à une crue, des données relatives à l'aléa, à l'occupation du sol et à la vulnérabilité des enjeux exposés au risque d'inondation sont croisées. On précise que la phase de modélisation du fonctionnement hydrologique et hydraulique du bassin versant et du cours d'eau n'est pas comprise dans le modèle ACB-DE, elle intervient uniquement en amont et est considérée comme une *variable d'entrée* du modèle.

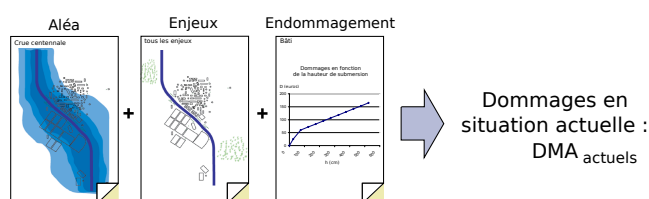
On définit les *Dommages Moyens Annualisés* (DMA) comme la somme des dommages dus aux différentes crues simulées, pondérés par la probabilité d'occurrence de chaque crue :

$$DMA = \int_{T_d}^{\infty} \frac{D(T)}{T} dT$$

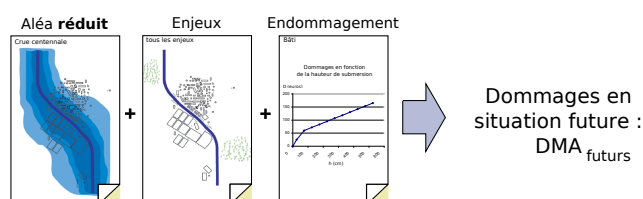
où $D(T)$ est le dommage (en euros) associé à la crue de période de retour T et T_d est la période de retour de la première crue qui crée des dommages (crue débordante). L'efficacité d'un aménagement de protection contre les inondations est évaluée en mesurant le différentiel entre deux situations : la situation sans aménagement (ou situation actuelle) et la situation avec aménagement (ou situation future). On définit ainsi les DEMA, indicateurs des bénéfices apportés par un aménagement (Figure 1) :

$$\text{DEMA} = \text{DMA}_{\text{actuel}} - \text{DMA}_{\text{futur}}$$

Situation sans aménagement de protection



Situation avec aménagement de protection



Bénéfices apportés par l'aménagement :
 $\text{DEMA} = \text{DMA}_{\text{actuels}} - \text{DMA}_{\text{futurs}}$

FIGURE 1. Principe de l'outil Analyse Coût-Bénéfice - Dommages Évités (ACB-DE).

Les données utilisées dans ce travail sont issues d'une étude de cas sur la basse vallée de l'Orb, dans l'Hérault [6]. La zone d'étude, d'une superficie de 100 km², englobe six communes à l'aval du bassin versant, de Béziers jusqu'à la mer (Figure 2). Ce secteur, qui accueille plus de 15 000 habitants, subit des crues fréquentes et constitue une zone à fort enjeu économique. En 2001, un schéma de protection contre les inondations a été élaboré par le Syndicat Béziers-la-Mer, qui proposait des aménagements de protection : des endiguements rapprochés, le rétablissement des exutoires en mer et l'amélioration de l'hydraulique dans la traversée de Béziers. L'effet de ces aménagements a été évalué au travers de l'outil ACB-DE.

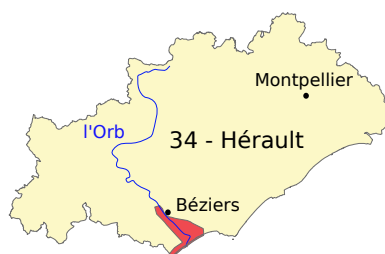


FIGURE 2. Zone d'étude : la basse vallée de l'Orb dans le département de l'Hérault. Superficie : 10000 ha.

3.2. Des variables d'entrée spatialisées

Le modèle ACB-DE nécessite six variables d'entrée, notées X_0 à X_5 (Table 1). Seul X_5 est une simple variable d'entrée scalaire. Deux variables d'entrée (X_2 et X_3) désignent chacune un *ensemble de variables scalaires*, qui sont rassemblées dans un même groupe pour les besoins de l'analyse de sensibilité (voir section 1.2.15 de Saltelli et al [19] pour une discussion sur l'analyse de sensibilité par groupes). Trois variables d'entrée (X_0 , X_1 et X_4) sont spatialisées. Le même terme *variable d'entrée* est utilisé pour désigner ces trois types d'entrée différentes (variable scalaire, groupe de variables scalaires, variable spatialisée). On précise que dans le cadre de ce papier, l'incertitude pesant sur les cartes de hauteurs de submersion (variable X_0) n'a pas été considérée : cet aspect, complexe, sera exploré dans de futurs travaux.

TABLEAU 1. Description des variables d'entrée de l'outil ACB-DE.

	Nom	Nature	Description	Unité
X_0	Altitudes d'eau (considéré invariant dans cette étude)	Variable spatialisée	Pour chaque crue simulée, deux cartes donnant les côtes maximales de la surface submergée, avec et sans projet de protection. Format : grilles 2D de résolution horizontale 5 m.	cm
X_1	Enjeux	Variable spatialisée	Carte permettant de localiser les enjeux (bâtiments, cultures, etc.) potentiellement exposés au risque d'inondation. Chaque enjeu est décrit par son type, sa surface et son seuil de protection. Format : couche SIG vectorielle.	Type : qualitatif Surface : m ² Seuil : cm
X_2	Périodes	Groupe de variables scalaires	Table donnant la période de retour de chaque crue simulée : crue débordante, crue décennale, crue trentennale, crue cinquantiennale et crue centennale.	années
X_3	Endommagement	Groupe de variables scalaires	Table donnant pour chaque type d'enjeu les coûts engendrés par une submersion en fonction de la hauteur de submersion.	€/m ²
X_4	MNT	Variable spatialisée	Modèle Numérique de Terrain donnant l'altitude du sol en tout point de la zone d'étude. Obtenu par stéréophotogrammétrie. Format : grille 2D de résolution horizontale 5 m.	m
X_5	C_∞	Variable scalaire	Coefficient multiplicatif utilisé pour évaluer les dommages dus à une crue extrême.	aucune

3.3. Une variable de sortie spatialisée

La variable de sortie considérée est l'indicateur DEMA. Cet indicateur peut être présenté sous deux formes : agrégé sur l'ensemble de la zone d'étude, c'est un *scalaire* (noté $\text{DEMA}^{(\text{tot})}$) qui indique les dommages évités totaux sur le territoire ; spatialisé, c'est alors une *carte* donnant en toute zone du territoire les dommages évités *sur cette zone*. Plus précisément, on note $\text{DEMA}^{(s)}[u]$ le champ spatial donnant l'indicateur DEMA calculé en toute cellule u d'une grille régulière de surface élémentaire s (Figure 3). L'indicateur DEMA est *spatialement additif*, en ce sens que l'indicateur DEMA défini sur l'union de deux zones disjointes est la somme des DEMA obtenus sur chacune des deux zones. On s'intéresse pour l'analyse de sensibilité à cinq variables de sortie différentes, qui sont cinq formes du même indicateur DEMA :

- $\text{DEMA}^{(\text{tot})}$: Dommages Évités Moyens Annualisés agrégés sur l'ensemble de l'aire d'étude (superficie de 10000 ha)
- $\text{DEMA}^{(4)}[u]$: carte des DEMA sur une grille régulière de surface élémentaire $s = 4$ ha
- $\text{DEMA}^{(16)}[u]$: carte des DEMA sur une grille régulière de surface élémentaire $s = 16$ ha
- $\text{DEMA}^{(64)}[u]$: carte des DEMA sur une grille régulière de surface élémentaire $s = 64$ ha
- $\text{DEMA}^{(256)}[u]$: carte des DEMA sur une grille régulière de surface élémentaire $s = 256$ ha

On précise que la résolution s des différentes cartes des DEMA ne constitue pas une variable d'entrée du modèle ACB-DE : il s'agit uniquement d'un choix de représentation spatiale de la sortie du modèle. Une même simulation du modèle ACB-DE permet de générer l'ensemble des cartes $\text{DEMA}^{(s)}[u]$ pour $s = 4, 16, 64$ et 256 ha.

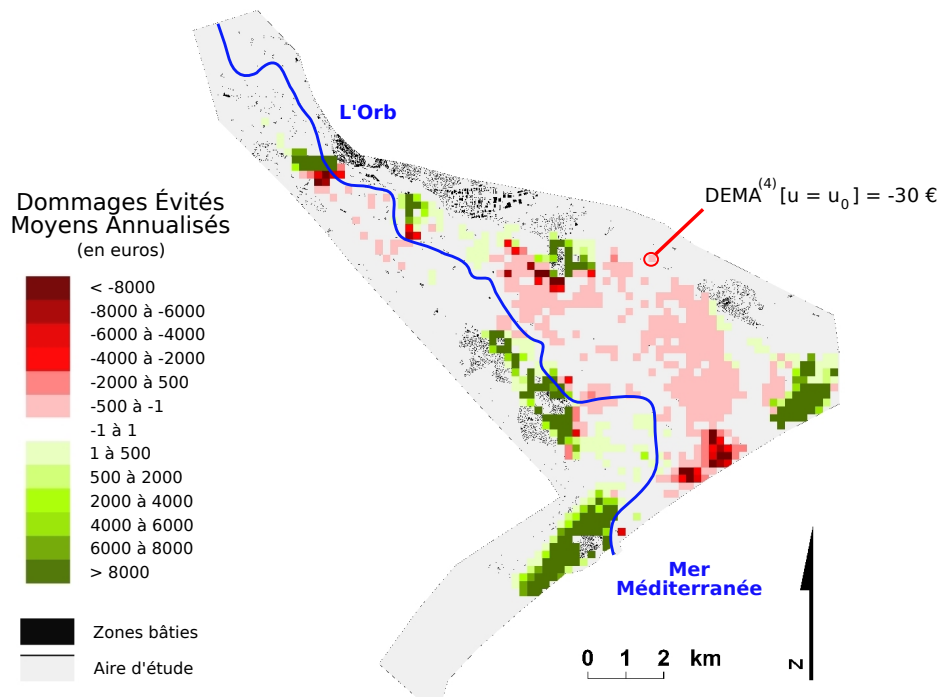


FIGURE 3. $\text{DEMA}^{(4)}[u]$: carte des Dommages Évités Moyens Annualisés sur une grille régulière de surface élémentaire $s = 4$ ha. En chaque cellule u_0 est représentée la valeur de l'indicateur DEMA calculé sur cette cellule, noté $\text{DEMA}^{(4)}[u = u_0]$. On considère ici que toutes les variables d'entrée prennent leur valeur nominale.

4. Analyse de sensibilité spatialisée du modèle ACB-DE : méthodes

L'approche proposée pour l'étude de l'outil ACB-DE est une analyse de sensibilité basée sur la variance, adaptée au cas d'un modèle dont les entrées et les sorties sont spatialisées. La mesure d'importance retenue est le couple (S_j, ST_j) des indices de Sobol de premier ordre et totaux. S_j représente la part de la variance de Y (sortie du modèle) qui serait supprimée si la variable X_j était connue avec certitude. ST_j représente la part résiduelle de la variance de Y si toutes les variables sauf X_j étaient connues avec certitude. Ces indices sont estimés à partir d'un échantillon aléatoire

de sorties du modèle, qui est simulé en un grand nombre de points de l'espace des variables d'entrée incertaines. On trouvera l'exposé des bases théoriques de l'analyse de sensibilité basée sur la variance ainsi que des précisions sur les différentes méthodes d'estimation des indices de Sobol dans Sobol' [21] et Saltelli et al [19].

L'analyse de sensibilité de l'outil ACB-DE s'est déroulée en trois étapes. Dans un premier temps, on a spécifié l'incertitude pesant sur chaque variable d'entrée. Afin de rendre compte de l'incertitude sur les deux variables spatialisées (X_1 , carte des enjeux et X_4 , Modèle Numérique de Terrain), deux jeux de $n_1 = 1000$ et $n_4 = 100$ réalisations aléatoires de ces variables ont été générés à partir de spécifications de l'incertitude spatialisée, prenant notamment en compte des effets de corrélation spatiale pour X_4 . Dans un second temps, un échantillon de taille $C = 28672$ a été constitué afin d'explorer l'espace des variables incertaines X_1 à X_5 , puis le modèle ACB-DE a été simulé en chaque point de cet échantillon. La variable spatialisée X_1 (resp. X_4) a été intégrée dans cet échantillon en associant un entier compris entre 1 et n_1 (resp. n_4) à chaque réalisation aléatoire préalablement générée de cette variable, ces réalisations étant considérées équiprobables. Dans un dernier temps, les indices de sensibilité de premier ordre et totaux de chaque variable d'entrée X_1 à X_5 ont été estimés pour la sortie scalaire $\text{DEMA}^{(\text{tot})}$. Pour les sorties spatialisées $\text{DEMA}^{(4)}[u]$, $\text{DEMA}^{(16)}[u]$, $\text{DEMA}^{(64)}[u]$ et $\text{DEMA}^{(256)}[u]$, des *cartes d'indices de sensibilité* ont été construites à différentes résolutions pour chaque variable d'entrée, afin de rendre compte de la variabilité spatiale des indices de sensibilité et de discuter de l'influence relative des variables d'entrée à plusieurs échelles.

4.1. Modélisation des incertitudes

Les sources d'incertitude qui affectent les variables d'entrée du modèle ACB-DE sont variées : erreurs de mesure, variabilité naturelle du phénomène représenté, manque de données et de connaissances, choix de modélisation (voir [15] pour une typologie des natures d'incertitude). On choisit ici de faire une description *probabiliste* de ces incertitudes et les variables d'entrée sont considérées indépendantes entre elles. Afin de rendre compte des incertitudes sur les variables spatialisées, il est fait appel à des modèles d'erreur, géostatistiques ou autres, qui permettent de générer un nombre fini de *réalisations aléatoires* de ces variables. La Table 2 et la Figure 4 présentent les spécifications de l'incertitude retenues pour chaque variable d'entrée.

4.1.1. Incertitude sur la carte des enjeux

La variable d'entrée X_1 est une carte *vectorielle* décrivant les enjeux exposés au risque d'inondation : chaque enjeu y est représenté par un polygone auquel sont associées trois caractéristiques (type, seuil de protection, surface), chacune étant affectée par une nature d'incertitude différente.

Type de l'enjeu le processus d'élaboration de la carte d'occupation du sol est entaché d'erreurs de photo-interprétation et de classification qui entraînent de possibles confusions entre les différents types d'enjeux. Pour rendre compte de cette incertitude, une *matrice de confusion* [8] donne pour chaque paire de types ($\text{type}_a, \text{type}_b$) une probabilité de confusion $p_{a,b}$. Cette matrice de confusion a été construite à dire d'expert, à partir de valeurs communément admises pour des données d'occupation du sol similaires.

TABLEAU 2. *Nature et spécification de l'incertitude sur les variables d'entrée du modèle ACB-DE.*

	Variable d'entrée	Nature des incertitudes	Spécification de l'incertitude et simulation
X_0	Altitudes d'eau	Incertainitudes non prises en compte dans le cadre de cette étude.	
X_1	Enjeux	Erreurs de classification	Matrice de confusion
		Variabilité des seuils de protection	Distribution empirique des seuils établie à partir d'un échantillon terrain
		Erreurs de mesure des surfaces	Application d'un coefficient multiplicatif de loi uniforme
X_2	Périodes	Manque de connaissance	Distribution uniforme entre T_{min} et T_{max} . Périodes de retour des différentes crues indépendantes.
X_3	Endommagement	Manque de connaissance	Pour chaque type d'enjeu, application d'un correctif multiplicatif aléatoire $(1 + \varepsilon_i)$. Coefficients ε_i indépendants, de loi uniforme dans $[-0.2; 0.2]$.
X_4	MNT	Erreurs de mesure et erreurs d'interpolation	Modélisation et simulation géostatistique.
X_5	C_∞	Choix de modélisation arbitraire	Distribution triangulaire symétrique (min = 1, max = 3).

Seuil de protection le seuil est la hauteur (en cm) de l'éventuelle surélévation du premier plancher qui protège un bâtiment face aux inondations. Suite à une enquête sur le terrain, cinq zones de bâti ont été identifiées sur l'aire d'étude : sur chacune de ces zones, la répartition statistique des seuils a été décrite par un histogramme empirique construit à partir d'un échantillon de bâtiments enquêtés.

Surface la surface des polygones décrivant les enjeux est utile au calcul des dommages, mais est entachée de multiples incertitudes. Pour en rendre compte, un correctif multiplicatif de surface, compris entre 0 et 1, est tiré dans une loi de probabilité propre à chaque type d'enjeu (Table 3).

À partir de cette description probabiliste de l'incertitude qui affecte chacun des trois champs de la carte des enjeux, un jeu de $n_1 = 1000$ réalisations aléatoires de cette carte a été généré. Le choix du nombre de cartes aléatoires générées répond à des contraintes de temps de calcul et

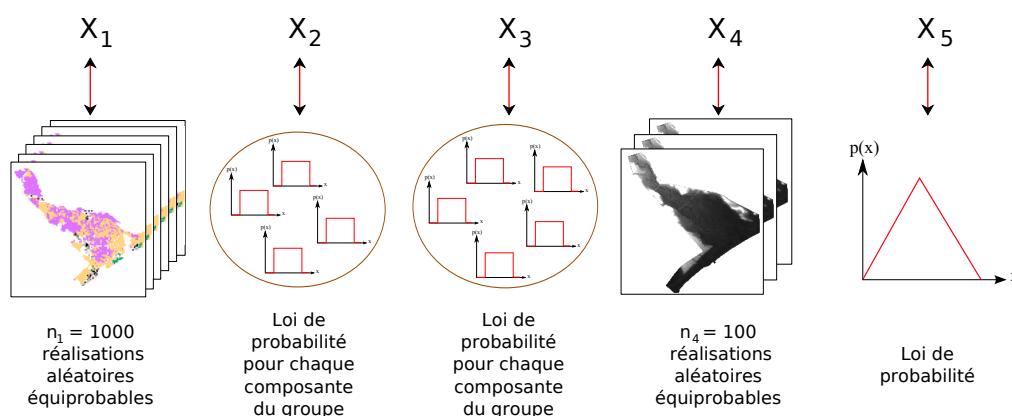


FIGURE 4. *Spécification de l'incertitude sur les variables d'entrée du modèle ACB-DE.*

d'espace de stockage disponible : la simulation d'une carte des enjeux coûte environ 5 secondes et le jeu de 1000 réalisations occupe 4 Go.

TABLEAU 3. *Loi de probabilité du correctif de surface des enjeux*

Type d'enjeu	Loi de probabilité du correctif de surface
Bâti	Uniforme dans $[0.75; 0.85]$
Activité	Uniforme $[0.75; 0.85] \times$ Uniforme $[0.4; 0.6]$
Culture	Fixe : 1
Camping	Fixe : 1

4.1.2. Incertitude sur le Modèle Numérique de Terrain

Le Modèle Numérique de Terrain (variable X_4) est construit à partir d'un jeu de 50 000 points de référence, répartis sur la zone d'étude, dont l'altitude a été estimée par stéréophotogrammétrie. L'altitude en tout point d'une grille régulière de résolution 5 m a ensuite été interpolée à partir de ce jeu de points de référence. Au moins deux types d'incertitude affectent le MNT : erreurs de mesure aux points de référence, erreurs d'interpolation ailleurs (Wechsler dresse un tableau complet des incertitudes associées aux Modèles Numériques de Terrain [23]). Ces incertitudes sont modélisées par un champ d'erreur aléatoire Gaussien ajouté au MNT initial, selon une méthode proposée par Castrignano [2]. Le MNT initial est ainsi considéré comme une représentation moyenne de la réalité, à laquelle est ajoutée un bruit aléatoire. Une campagne terrain a permis de déterminer les caractéristiques de ce champ d'erreur : des altitudes d'une précision centimétrique ont été relevées par GPS différentiel en 500 points de contrôle répartis sur l'aire d'étude. Sur la base de ces données, la structure spatiale du champ d'erreur aléatoire Δ a été décrite par un modèle de variogramme exponentiel (portée effective de 500 m, effet de pépité $\eta = 0.02$, palier de 0.11 m^2). Un jeu de $n_4 = 100$ réalisations aléatoires du champ d'erreur Δ a ensuite été généré selon un algorithme de simulation gaussienne séquentielle [5], conditionnellement aux erreurs relevées aux points de contrôle (outil SGSIM du logiciel SGeMS [1]). Le choix de n_4 a été guidé par des considérations de coût en temps de calcul, chaque simulation durant environ 2 minutes. Les champs d'erreur ainsi simulés vérifient les conditions suivantes :

- reproduction des valeurs de l'erreur Δ aux points de contrôle ;
- reproduction de la distribution empirique de l'erreur Δ observée sur les points de contrôle ;
- reproduction (en moyenne sur l'ensemble des 100 simulations) du modèle de variogramme du champ d'erreur Δ .

Ces champs d'erreur ont ensuite été ajoutés au Modèle Numérique de Terrain initial, pour obtenir un jeu de $n_4 = 100$ MNT bruités, qui sont considérés comme 100 réalisations aléatoires équiprobables de la variable d'entrée X_4 (Figure 5).

4.1.3. Incertitude sur les variables non spatialisées

La variable d'entrée X_2 est un groupe de cinq variables scalaires. Ces variables sont les périodes de retour T de chaque crue étudiée (débordante, décennale, trentennale, cinquantennale, centennale).

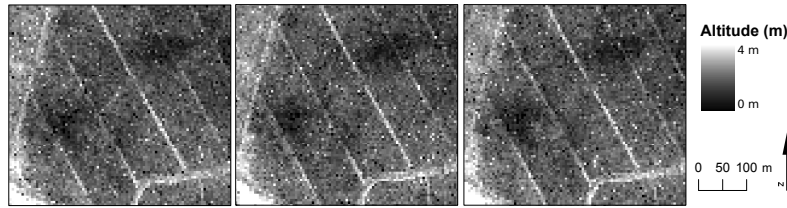


FIGURE 5. *Trois Modèles Numériques de Terrain bruités par addition d'un champ d'erreur Gaussien. Zoom sur une zone de 18 ha.*

Elles sont supposées indépendantes et sont décrites par des lois uniformes dont les bornes T_{\min} et T_{\max} sont données par l'intervalle de confiance à 95 % de l'ajustement de la loi de Gumbel (régression non linéaire) sur la série de débits instantanés de la station hydrométrique de Tabarka (observés de 1967 à 2008), station sur l'Orb en entrée de la zone d'étude (Table 4).

La variable d'entrée X_3 est un groupe de courbes d'endommagement, au nombre de 45 (une courbe par type d'enjeu). L'incertitude sur chacune de ces courbes est décrite par l'application d'un correctif multiplicatif aléatoire $(1 + \varepsilon_i)_{1 \leq i \leq 45}$ où les ε_i sont indépendants et suivent une loi uniforme dans $[-0.2; 0.2]$.

Enfin, la variable d'entrée X_5 (C_∞) est supposée suivre une loi de probabilité triangulaire de minimum 1 et de maximum 3.

TABEAU 4. *Incertaince sur la variable d'entrée X_2 : périodes de retour des crues*

Crue simulée	Débit ($\text{m}^3 \cdot \text{s}^{-1}$)	T_{\min} (années)	T_{\max} (années)
Débordante	non précisé	3	7
Décennale	929	9.3	10.7
Trentennale	1223	27.0	33.4
Quintennale	1357	44.2	56.6
Centennale	1538	86.3	116.0

4.2. Propagation des incertitudes

L'estimation des indices de sensibilité de Sobol s'appuie sur un échantillonnage intensif dans l'espace des variables d'entrée incertaines. L'approche utilisée est celle proposée par Lilburne et Tarantola [13], qui adapte les méthodes de Sobol et Saltelli [21, 19] au cas de variables d'entrée spatialisées. Deux échantillons aléatoires A et B de même dimension $p \times N$ sont tirés ($p = 5$ est le nombre de variables d'entrée, $N = 4096$ la taille des échantillons). La $i^{\text{ème}}$ ligne d'un échantillon (A ou B) est un jeu $(X_1^{(i)}, X_2^{(i)}, X_3^{(i)}, X_4^{(i)}, X_5^{(i)})$ de variables d'entrée, pour lequel le modèle ACB-DE peut être simulé. La $j^{\text{ème}}$ colonne de A ou de B est un échantillon de réalisations aléatoires de la variable d'entrée incertaine X_j . Pour la variable scalaire X_5 , chaque élément $X_5^{(i)}$ est un scalaire, tiré dans la loi de probabilité de X_5 . Pour les variables d'entrée X_2 et X_3 , qui sont des groupes de variables scalaires, chaque élément $X_2^{(i)}$ et $X_3^{(i)}$ est un *vecteur* dont les composantes sont tirées indépendamment dans leurs lois de probabilité respectives.

Prise en compte des variables d'entrée spatialisées Pour intégrer la variable spatialisée X_1 (carte des enjeux) à cet échantillonnage, on considère que la variabilité de X_1 est représentée par le jeu de $n_1 = 1000$ réalisations aléatoires de la carte des enjeux qui a été préalablement généré. On considère que ces n_1 réalisations sont *équiprobables* et l'on associe à chacune d'elles un entier compris entre 1 et n_1 . Chaque élément $X_1^{(i)}$ des échantillons A et B est alors un entier tiré au hasard entre 1 et n_1 : c'est une réalisation d'une variable aléatoire discrète de loi uniforme dans l'ensemble $\llbracket 1, n_1 \rrbracket$. La valeur de $X_1^{(i)}$ indique le *numéro de la réalisation aléatoire* de la carte des enjeux à utiliser pour simuler le modèle pour cette ligne de l'échantillon. Le même principe est utilisé pour la variable spatialisée X_4 (Modèle Numérique de Terrain), pour lequel $n_4 = 100$ réalisations aléatoires avaient été préalablement générées.

Cette procédure induit un biais dans la prise en compte de l'incertitude sur les variables d'entrée spatialisées X_1 et X_4 : leur variabilité totale est représentée par un nombre fini de réalisations aléatoires. Cette limitation est due à des considérations pratiques, notamment aux coûts liés à la simulation et au stockage d'un très grand nombre de cartes. On peut cependant noter que dès lors que le nombre n de réalisations aléatoires pré-générées est plus grand que la taille N des échantillons A et B , le biais dû au pré-échantillonnage disparaît, à condition de tirer aléatoirement et *sans remise* parmi le jeu de cartes pré-générées.

Permutations Afin de pouvoir estimer les espérances des variances conditionnelles qui définissent les indices de sensibilité, le modèle ACB-DE doit être simulé en des points où seule l'une des variables X_j varie par rapport à un point $(X_1^{(i)}, X_2^{(i)}, X_3^{(i)}, X_4^{(i)}, X_5^{(i)})$ où le modèle a déjà été simulé. À cette fin, de nouveaux échantillons sont créés par permutations entre les échantillons A et B . Pour j allant de 1 à $p = 5$, on crée une matrice $A_B^{(j)}$, identique à la matrice A , sauf dans sa $j^{\text{ème}}$ colonne qui est celle de B (Figure 6, voir [18] pour les détails de cette procédure).

$$A = \begin{pmatrix} X_1^{(1)} & \dots & X_j^{(1)} & \dots & X_5^{(1)} \\ \vdots & & \ddots & & \vdots \\ X_1^{(N)} & \dots & X_j^{(N)} & \dots & X_5^{(N)} \end{pmatrix} \quad B = \begin{pmatrix} X_1^{(N+1)} & \dots & X_j^{(N+1)} & \dots & X_5^{(N+1)} \\ \vdots & & \ddots & & \vdots \\ X_1^{(2N)} & \dots & X_j^{(2N)} & \dots & X_5^{(2N)} \end{pmatrix}$$

$$A_B^{(j)} = \begin{pmatrix} X_1^{(1)} & \dots & X_j^{(N+1)} & \dots & X_5^{(1)} \\ \vdots & & \ddots & & \vdots \\ X_1^{(N)} & \dots & X_j^{(2N)} & \dots & X_5^{(N)} \end{pmatrix}$$

FIGURE 6. Création de l'échantillon $A_B^{(j)}$

Simulations du modèle On simule ensuite le modèle ACB-DE pour chacune des lignes des échantillons A , B , $A_B^{(j)}$ (Figure 7). Le nombre total de simulations du modèle est $C = (p + 2) \times N$, soit $C = 28672$. Chaque simulation du modèle dure environ 5 secondes : le coût total de cette analyse de sensibilité a été de 40 heures de calcul, réparties sur deux processeurs.

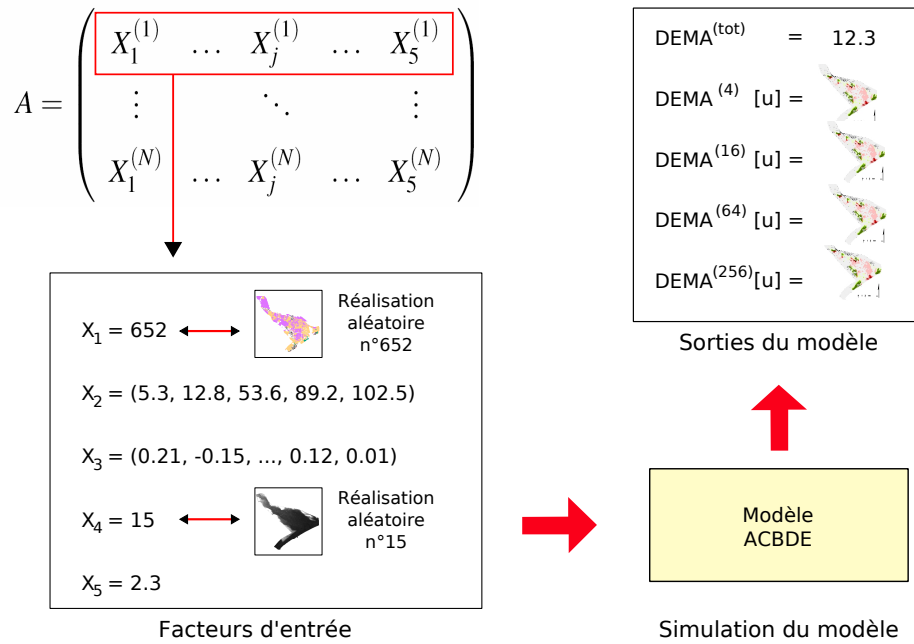


FIGURE 7. Simulation du modèle pour une ligne d'échantillon

4.3. Estimation des indices de sensibilité et cartes d'indices de sensibilité

La simulation du modèle ACB-DE pour une ligne i d'échantillon permet d'obtenir une valeur pour les cinq sorties étudiées : $DEMA^{(tot)}$ (sortie scalaire) et $DEMA^{(4)}[u]$ à $DEMA^{(256)}[u]$ (sorties spatialisées sur des grilles de résolution s décroissante). À chaque sortie spatialisée $DEMA^{(s)}[u]$ ($s = 4$ à 256 ha) correspond une liste de sorties scalaires $(DEMA^{(s)}[u = u_0])_{u_0 \in \mathbb{G}^{(s)}}$, qui donnent l'indicateur DEMa calculé en chaque cellule u_0 de la grille $\mathbb{G}^{(s)}$ de résolution s .

On note de manière générique Y l'une quelconque de ces sorties scalaires ($DEMA^{(tot)}$ ou l'un des $DEMA^{(s)}[u = u_0]$). L'évaluation de Y pour chacune des lignes des échantillons A, B et $A_B^{(j)}$, permet d'obtenir trois vecteurs de taille N , que l'on note Y_A, Y_B et $Y_{A_B^{(j)}}$. On estime alors les indices de sensibilité de premier ordre et totaux de chacune des variables X_j par rapport à Y , notés $S_j(Y)$ et $ST_j(Y)$, selon les expressions (1) et (2) recommandées par Saltelli et al. [18]. Des intervalles de confiance à 90 % sont estimés pour chacun des indices de sensibilité par bootstrap (nombre d'échantillons bootstrap $n_{boot} = 100$).

$$S_j = \frac{\frac{1}{N} \sum_{i=1}^N Y_B^{(i)} \cdot Y_{A_B^{(j)}}^{(i)} - \frac{1}{N} \sum_{i=1}^N Y_B^{(i)} \cdot Y_A^{(i)}}{\frac{1}{N} \sum_{i=1}^N Y_A^{(i)} \cdot Y_A^{(i)} - \left(\frac{1}{N} \sum_{i=1}^N Y_A^{(i)} \right) \cdot \left(\frac{1}{N} \sum_{i=1}^N Y_B^{(i)} \right)} \quad (1)$$

$$ST_j = \frac{\frac{1}{2N} \sum_{i=1}^N \left[Y_A^{(i)} - Y_{A_B}^{(i)} \right]^2}{\frac{1}{N} \sum_{i=1}^N Y_A^{(i)} \cdot Y_A^{(i)} - \left(\frac{1}{N} \sum_{i=1}^N Y_A^{(i)} \right) \cdot \left(\frac{1}{N} \sum_{i=1}^N Y_B^{(i)} \right)} \quad (2)$$

Réalisation de cartes d'indices de sensibilité Pour chacune des sorties spatialisées $DEMA^{(4)}[u]$ à $DEMA^{(256)}[u]$, on établit des *cartes de sensibilité* pour chaque variable d'entrée X_1 à X_5 . Ces cartes sont définies sur la même grille régulière $\mathbb{G}^{(s)}$ que la sortie $DEMA^{(s)}[u]$ considérée. La valeur de la carte de sensibilité de la variable X_j en une cellule u_0 de la grille $\mathbb{G}^{(s)}$ est égale à l'indice de sensibilité total $ST_j(DEMA^{(s)}[u = u_0])$ de cette variable par rapport à la sortie scalaire $DEMA^{(s)}[u = u_0]$ définie sur cette cellule. On obtient ainsi un jeu de 20 cartes de sensibilité qui traduisent l'influence de l'incertitude des différentes variables d'entrée X_1 à X_5 sur la variabilité de l'indicateur DEMA à différentes échelles d'étude. On a choisi de ne considérer que les cartes d'indices de sensibilité totaux : en effet, les intervalles de confiance à 90 % sur les indices de sensibilité, calculés par bootstrap, sont ici plus étroits pour les indices totaux que pour les indices de premier ordre, pour la taille N d'échantillon retenue (cette observation empirique n'a pour l'instant pas pu être expliquée).

5. Résultats

Variabilité de l'indicateur $DEMA^{(tot)}$ La Figure 8 montre la distribution empirique de l'indicateur $DEMA^{(tot)}$, obtenue sur l'ensemble des $C = 28\,672$ simulations du modèle ACB-DE. La moyenne empirique est $\mu_{DEMA^{(tot)}} = 8.63$ M€, peu éloignée de la valeur nominale $DEMA_{nom}^{(tot)} = 8.31$ M€, obtenue pour une simulation du modèle où chaque variable d'entrée prend sa valeur nominale. Cette valeur de Dommages Évités Moyens Annualisés est à comparer aux coûts d'investissement initiaux nécessaires à la réalisation des aménagements de protection réalisés sur l'Orb (32 M€), ainsi qu'à leurs coûts annuels d'entretien (1 M€) : ces aménagements s'avèrent économiquement rentables au bout de cinq ans. L'écart-type empirique $\sigma = 1.23$ M€ montre que la variabilité de l'indicateur $DEMA^{(tot)}$ est forte au regard de la précision raisonnablement attendue par les décideurs. Cependant, $DEMA^{(tot)}$ ne prend sur l'ensemble des simulations que des valeurs positives. S'il est difficile de chiffrer avec une précision convenable les dommages évités sur l'ensemble du territoire, il semble en revanche possible de conclure que les aménagements de protection mis en place sur l'Orb ont un impact positif, puisque qu'ils permettent de réduire les dommages totaux sur le territoire.

Indices de sensibilité par rapport à la sortie scalaire $DEMA^{(tot)}$ La Figure 9 montre les indices de sensibilité (de premier ordre et totaux) de chacune des cinq variables d'entrée X_1 à X_5 par rapport à la sortie scalaire $DEMA^{(tot)}$. Il apparaît une démarcation claire entre deux groupes de variables. D'un côté, l'incertitude pesant sur les variables Périodes (X_2) et Endommagement (X_3) explique plus de 75 % de la variance de $DEMA^{(tot)}$: $S_2 + S_3 = 0.7625$. De l'autre côté, le caractère incertain des variables Enjeux (X_1), MNT (X_4) et C_∞ (X_5) n'explique qu'une très faible part de la variabilité de $DEMA^{(tot)}$: $S_1 + S_4 + S_5 < 0.05$. La comparaison des indices de premier ordre S_j et des indices totaux ST_j indique de plus que les interactions des variables Enjeux (X_1),

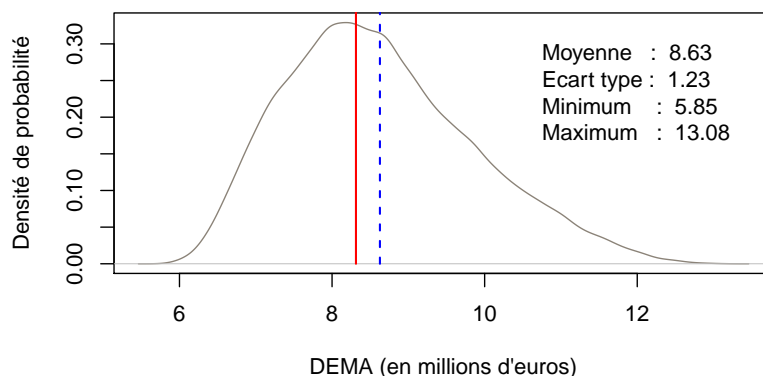


FIGURE 8. Densité de probabilité empirique de l'indicateur $DEMA^{(tot)}$. Estimation par noyau Gaussien. Nombre de simulations : $N = 28672$. Taille de la fenêtre de lissage : $h = 0.128$ M€. Le trait pointillé vertical indique la moyenne empirique, le trait plein la valeur de l'indicateur lorsque toutes les variables d'entrée prennent leur valeur nominale.

MNT (X_4) et C_∞ (X_5) avec les autres variables ne contribuent pas à la variabilité de $DEMA^{(tot)}$. Seule l'interaction entre les variables Périodes (X_2) et Endommagement (X_3) explique une part significative de la variance de l'indicateur : $S_{2,3} \approx ST_2 - S_2 \approx ST_3 - S_3 \approx 0.12$. Dans une optique de réduction de la variance de l'indicateur $DEMA^{(tot)}$, il apparaît donc indispensable d'améliorer la précision des deux variables X_2 et X_3 , tandis qu'aucun effort supplémentaire n'a besoin d'être fait sur les variables X_1 , X_4 et X_5 .

Cartes d'indices de sensibilité La Figure 10 présente les *cartes d'indices de sensibilité* des cinq variables d'entrée X_1 à X_5 par rapport à deux sorties spatialisées : $DEMA^{(4)}[u]$ et $DEMA^{(256)}[u]$. Sur chacune de ces cartes, on observe une variabilité spatiale des indices de sensibilité. À une résolution donnée, l'influence d'une variable incertaine sur la variabilité de l'indicateur DEMA calculé localement (en une cellule de la grille) varie d'une zone à l'autre du territoire. On peut

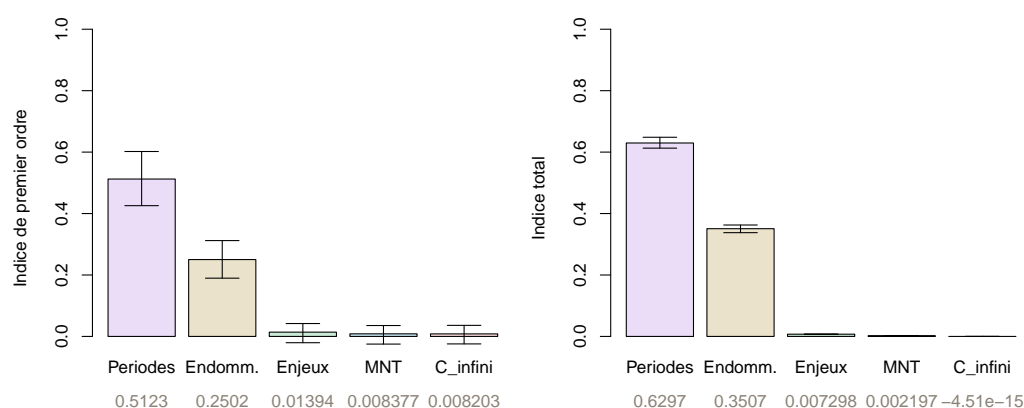


FIGURE 9. Indices de sensibilité de premier ordre S_j (gauche) et totaux ST_j (droite) des cinq variables d'entrée du modèle ACB-DE par rapport à la variable $DEMA^{(tot)}$ et leurs intervalles de confiance à 90 % estimés par bootstrap.

identifier deux éléments de structuration spatiale majeurs : les effets de bord et la présence de villes (bâti urbain). Sur les marges du territoire étudié, les indices de sensibilité des variables MNT (X_4) et Enjeux (X_1) sont plus faibles qu'au coeur de la vallée. Dans les zones urbaines (identifiées sur la Figure 10 par des limites pointillées), les indices de sensibilité des variables Périodes (X_2) et Endommagement (X_3) sont plus importants qu'en zone non urbaine, tandis que les indices des variables MNT et Enjeux sont plus faibles qu'en zone non urbaine.

Par ailleurs, on observe que les cartes de sensibilité par rapport à la sortie $\text{DEMA}^{(4)}[u]$ sont différentes des cartes réalisées par rapport à la sortie $\text{DEMA}^{(256)}[u]$. Pour les variables MNT (X_4) et Enjeux (X_1), la carte de sensibilité par rapport à $\text{DEMA}^{(256)}[u]$ présente des valeurs globalement moins élevées que la carte de sensibilité par rapport à $\text{DEMA}^{(4)}[u]$. Au contraire, pour les variables Périodes (X_2) et Endommagement (X_3), la carte de sensibilité par rapport à $\text{DEMA}^{(256)}[u]$ présente des valeurs globalement plus élevées que la carte de sensibilité par rapport à $\text{DEMA}^{(4)}[u]$. Pour la variable C_∞ (X_5), les cartes aux deux résolutions présentent des valeurs d'indices similaires, toutes inférieures à 0.1.

Pour permettre une comparaison entre les cartes d'indices de sensibilité aux différentes échelles, on a calculé pour chaque variable d'entrée X_j et chaque résolution s ($s = 4, 16, 64$ et 256 ha) la valeur moyenne de l'indice de sensibilité total $ST_j(\text{DEMA}^{(s)}[u = u_0])$ sur l'ensemble des cellules u_0 de la grille $\mathbb{G}^{(s)}$. On note $\mu_j^{(s)}$ cet *indice de sensibilité moyen* (3) :

$$\mu_j^{(s)} = \frac{1}{\text{card}(\{u_0 \in \mathbb{G}^{(s)}\})} \cdot \sum_{u_0 \in \mathbb{G}^{(s)}} ST_j(\text{DEMA}^{(s)}[u = u_0]) \quad (3)$$

La Figure 11 montre que pour chaque variable d'entrée X_j , l'indice moyen $\mu_j^{(s)}$ dépend de manière monotone de la résolution s . Pour les variables MNT (X_4) et Enjeux (X_1), $\mu_j^{(s)}$ est une fonction décroissante de la surface élémentaire s : plus l'indicateur DEMa est étudié à des résolutions grossières (s grand), moins sa variabilité est expliquée par l'incertitude sur la carte des enjeux ou le Modèle Numérique de Terrain. Au contraire, pour les variables Périodes (X_2) et Endommagement (X_3), $\mu_j^{(s)}$ est une fonction croissante de la surface élémentaire s : plus l'indicateur DEMa est étudié à des résolutions grossières (s grand), plus est élevée la part de sa variabilité qui est expliquée par l'incertitude sur les périodes de retour et les courbes d'endommagement. Il apparaît ainsi que *l'échelle d'étude conditionne la hiérarchie des variables incertaines* : à grande échelle (s petit), les variables présentant les plus forts indices de sensibilité sont MNT (X_4) et Enjeux (X_1), alors qu'à petite échelle (territoire pris dans son ensemble), les variables dont les indices de sensibilité sont les plus élevés sont Périodes (X_2) et Endommagement (X_3).

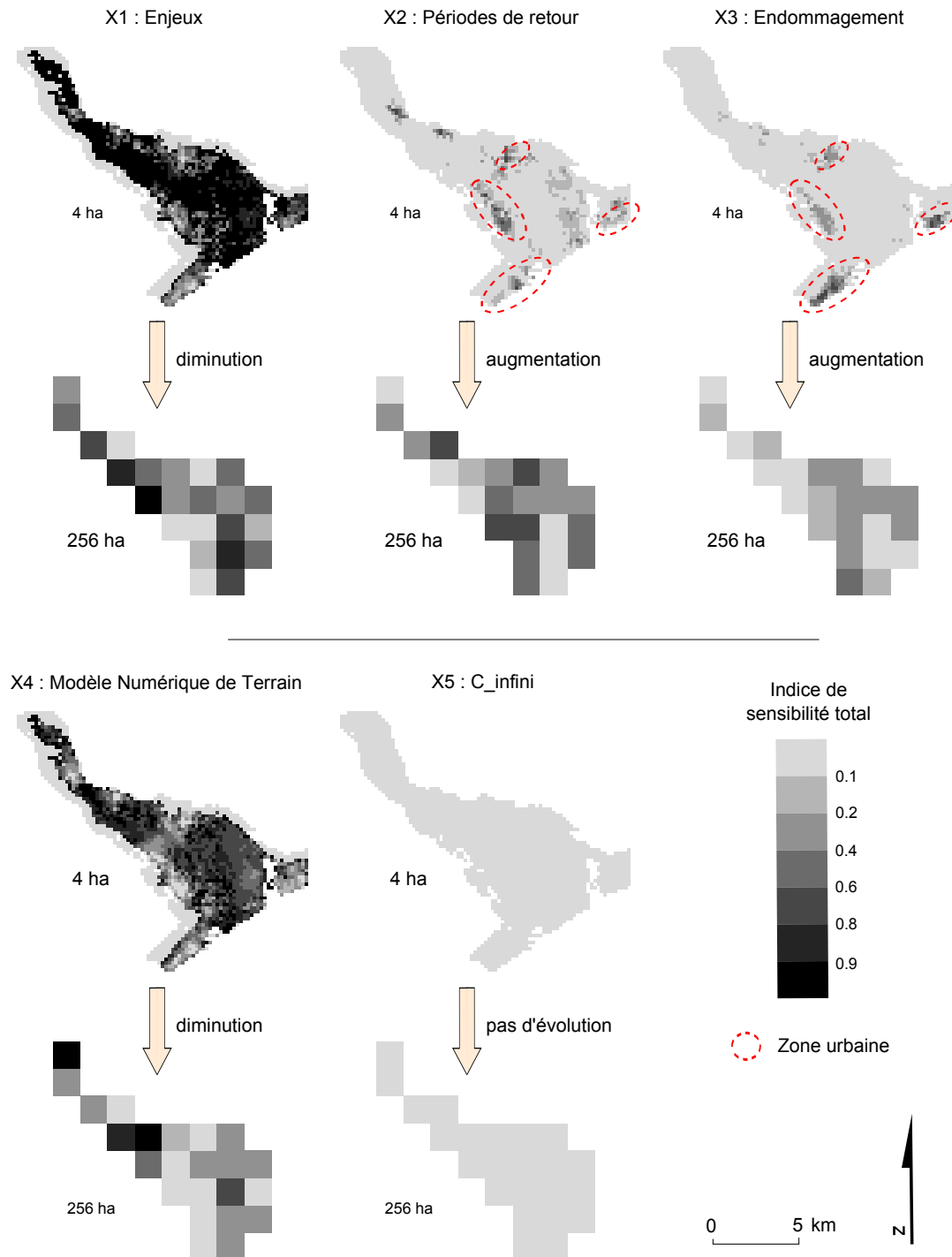


FIGURE 10. Cartes de sensibilité. Pour chacun des cinq variables X_1 à X_5 , la carte du haut (resp. du bas) présente la « carte de sensibilité » de la variable X_j par rapport à la sortie spatialisée $DEMA^{(4)}[u]$ (resp. $DEMA^{(256)}[u]$) : sur chaque cellule u_0 de la grille régulière de surface $s = 4$ ha (resp. 256 ha), on représente $ST_j(DEMA^{(4)}[u = u_0])$ (resp. $ST_j(DEMA^{(256)}[u = u_0])$), indice de sensibilité total de la variable X_j par rapport à la sortie scalaire $DEMA^{(4)}[u = u_0]$ (resp. $DEMA^{(256)}[u = u_0]$).

Soumis au Journal de la Société Française de Statistique

File: papierSFds_final.tex, compiled with jsfds, version : 2009/12/09

date: 15 février 2011

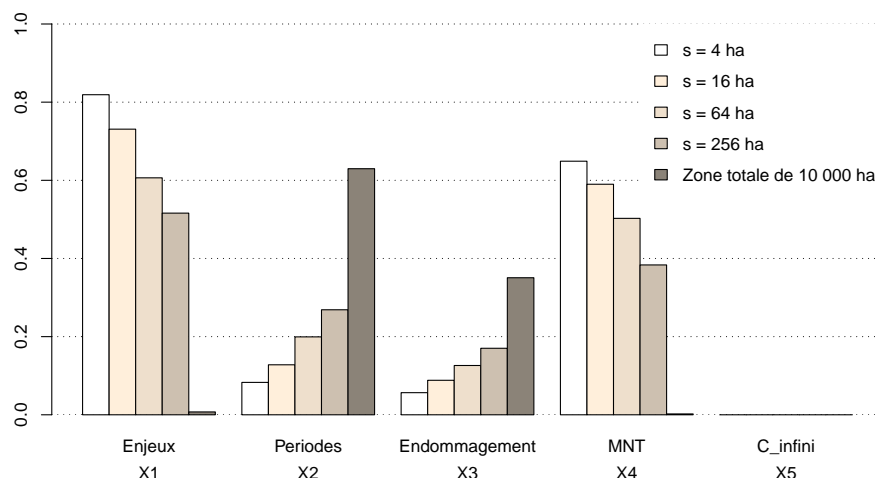


FIGURE 11. Indices de sensibilité moyens $\mu_j^{(s)}$. Pour quatre grilles de surface élémentaire s croissante et pour chaque variable d'entrée X_j , $\mu_j^{(s)}$ indique la valeur moyenne de la carte de sensibilité de X_j par rapport à $DEMA^{(s)}[u]$. Pour chaque variable X_j , la dernière barre indique la valeur de $ST_j(DEMA^{(tot)})$ (indice de sensibilité total de X_j par rapport à la sortie scalaire $DEMA^{(tot)}$).

6. Discussion

Ce travail visait à explorer le potentiel de l'analyse de sensibilité globale basée sur la variance pour étudier un modèle peu coûteux en temps de calcul et dont les entrées et les sorties sont spatialisées. L'étude du modèle ACB-DE confirme l'intérêt d'une telle approche : l'analyse menée a conduit à une hiérarchisation des variables d'entrée spatialisées, sur lesquelles des spécifications complexes de l'incertitude avaient été définies ; elle a mis en évidence la variabilité spatiale des indices de sensibilité ; elle a apporté un éclairage nouveau sur les liens entre la sensibilité des différentes variables d'entrée et l'échelle d'étude, échelle à laquelle on utilise les résultats du modèle ACB-DE.

Intégrer des variables d'entrée spatialisées à l'analyse de sensibilité La méthode proposée a permis de spécifier des incertitudes spatialisées pour les deux variables X_1 (carte des enjeux) et X_4 (Modèle Numérique de Terrain) et de les intégrer toutes deux à l'analyse de sensibilité aux côtés de variables scalaires. Les résultats montrent cependant que ces deux variables n'ont qu'une influence négligeable sur la variabilité de l'indicateur $DEMA^{(tot)}$, la plus grande part de la variabilité étant due à l'incertitude sur deux variables non spatialisées : Périodes (X_2) et Endommagement (X_3). L'approche retenue n'est pas limitée à l'intégration de variables d'entrée spatialisées, mais pourrait être appliquée à tout type de variable d'entrée sur laquelle il est nécessaire de spécifier l'incertitude de manière complexe : entrée fonctionnelle, temporelle, etc... Le principal défaut de cette approche réside dans le *pré-échantillonnage* qui est effectué pour les variables d'entrée spatialisées. L'incertitude pesant sur ces cartes est artificiellement représentée par un nombre fini n de réalisations aléatoires, considérées équiprobables. Le choix de n est contraint par des considérations d'ordre technique, telles que le temps de calcul nécessaire à la

simulation d'une réalisation aléatoire, ou l'espace disque nécessaire à son stockage. Le biais introduit par cet artifice diminue avec n . Dès que n est supérieur à la taille N des échantillons utilisés pour estimer les indices de Sobol, le biais peut être considéré comme nul (à condition de tirer aléatoirement les cartes pré-générées *sans remise*).

Les autres limites ne sont pas spécifiques à l'approche décrite dans ce papier, mais sont communes aux études basées sur l'estimation des indices de Sobol par des procédures d'échantillonnage de Monte-Carlo : le coût en nombre de simulations est très élevé, potentiellement rédhibitoire pour certains modèles, les variables d'entrée X_j doivent être indépendantes entre elles et la variabilité de la sortie est résumée à son seul moment du second ordre. Enfin, il ne faut pas oublier que l'incertitude sur les variables d'entrée, dont dépendent les résultats de l'analyse de sensibilité, est souvent difficile à spécifier par manque de données.

Cartes d'indices de sensibilité pour étudier la variabilité spatiale des indices de Sobol La réalisation de *cartes de sensibilité* permet de mobiliser les indices de Sobol pour étudier une sortie de modèle spatialisée. Ces cartes de sensibilité ont mis en évidence la variabilité spatiale des indices des différentes variables d'entrée par rapport aux sorties $\text{DEMA}^{(4)}[u]$, $\text{DEMA}^{(16)}[u]$, $\text{DEMA}^{(64)}[u]$ et $\text{DEMA}^{(256)}[u]$: les influences relatives des cinq variables d'entrée diffèrent selon que l'on se trouve en zone urbaine ou non, ou bien sur les marges du territoire étudié. On peut par exemple en conclure que l'incertitude affectant les courbes d'endommagement (X_3) a un plus fort impact sur la précision de l'indicateur DEMA en zones urbaines (identifiées sur la Figure 10 par des limites pointillées) qu'en zone non urbanisée. Ceci peut s'expliquer par le fait que les valeurs d'endommagement pour les enjeux de type *bâti* ou *activités économiques*, situés en zones urbaines, sont plus fortes que les valeurs d'endommagement pour les enjeux de type *cultures*, situés en zones non urbaines : l'introduction d'une même incertitude relative de $\pm 20\%$ sur toutes les courbes d'endommagement se traduit par une incertitude en valeur absolue plus forte en zone urbaine, et donc par un indice de sensibilité plus grand de la variable X_3 . Ces informations sont précieuses et permettront de diriger les futurs efforts de recherche pour l'amélioration de l'outil ACB-DE. Les cartes de sensibilité apparaissent donc comme une nouvelle source d'information, qui reste à exploiter pleinement.

Lien entre sensibilité des variables d'entrée et échelle d'étude La réalisation de cartes de sensibilité à différentes résolutions a mis en évidence un lien fort entre les indices de sensibilité des variables d'entrée et l'échelle d'étude, échelle à laquelle on analyse les sorties du modèle ACB-DE. À petite échelle, c'est-à-dire lorsque l'on considère la zone d'étude dans son ensemble, deux variables expliquent la plus grande part de la variance de l'indicateur DEMA : Périodes (X_2) et Endommagement (X_3). Ces deux variables sont des groupes de variables scalaires, qui ne sont pas distribuées spatialement. À grande échelle, c'est-à-dire lorsqu'on évalue l'indicateur DEMA sur des cellules de surface $s = 4$ ha, X_2 et X_3 jouent un rôle négligable et ce sont les variables Enjeux (X_1) et MNT (X_4) qui contribuent majoritairement à la variance de l'indicateur DEMA. Ces deux variables sont elles distribuées spatialement.

Le calcul des *indices moyens* $\mu_j^{(s)}$ à différentes résolutions apporte une interprétation plus générale : l'indice de sensibilité des variables *distribuées spatialement* (ici X_1 et X_4) diminue à mesure qu'augmente la taille s de la zone élémentaire sur laquelle on observe l'indicateur DEMA. Un effet de compensation des incertitudes en est responsable : les incertitudes qui affectent les

variables spatialisées sont *locales* ; lorsqu'une large zone est considérée, ces incertitudes locales se compensent et leur impact relatif par rapport aux autres variables diminue. Symétriquement, l'indice de sensibilité des variables *non distribuées spatialement* (ici X_2 , X_3 et X_5) augmente avec la taille s de la zone élémentaire sur laquelle on observe l'indicateur DEMA. Selon l'échelle à laquelle on observe l'indicateur DEMA, les variables d'entrée qui expliquent sa variabilité ne sont donc pas les mêmes.

Ces conclusions ne sont pas spécifiques à l'outil ACB-DE : les effets d'échelles mis en évidence sont communs à tous les modèles spatialisés. L'analyse de sensibilité spatialisée présentée ici constitue un nouvel outil pour explorer quantitativement ces effets liés aux changements d'échelle ; elle peut apporter des réponses à la question « *Pour utiliser ce modèle à telle résolution, sur quelles variables d'entrée dois-je porter mon attention en priorité ?* ». On notera cependant que cette discussion n'est valable que dans le cas d'une sortie spatialisée *additive*, telle que l'indicateur DEMA.

Perspectives Plusieurs questions doivent maintenant être explorées pour conforter la méthode présentée dans ce papier. Le biais introduit dans l'estimation des indices de sensibilité par le *pré-échantillonnage* des variables spatialisées devra être quantifié, afin de déterminer la taille n minimale du jeu de réalisations aléatoires à générer. Pour un n fixé, le choix d'un échantillonnage *optimal* des réalisations aléatoires de cartes est aussi problématique [17] ; des approches telles que celle proposée par Scheidt et al. [20], basée sur une mesure de la « *dissimilarité* » entre des réalisations géostatistiques, seront étudiées. Enfin, il sera intéressant d'explorer l'influence de la corrélation spatiale interne à chaque variable d'entrée spatialisée, ainsi que l'influence de la résolution spatiale de ces variables d'entrée, sur la valeur de leurs indices de sensibilité. Pour traiter ces questions, on étudiera un modèle spatialisé semblable à l'outil ACB-DE, à l'expression analytique simple.

7. Conclusion

On a décrit dans ce papier une application de l'analyse de sensibilité globale basée sur la variance à l'outil spatialisé ACB-DE, qui est un modèle d'évaluation économique du risque d'inondation, utilisé ici sur la basse-vallée de l'Orb (Hérault). L'approche proposée a permis (1) d'intégrer à l'analyse des variables d'entrée spatialisées en spécifiant de manière sophistiquée l'incertitude les affectant (2) d'étudier les sorties spatialisées du modèle ACB-DE par la réalisation de cartes d'indices de sensibilité, qui rendent compte de la variabilité spatiale des indices de Sobol (3) de comparer l'influence relative des variables d'entrée à différentes échelles, par la réalisation de cartes de sensibilité de résolution croissante. Il apparaît que les variables d'entrée spatialisées incertaines (Modèle Numérique de Terrain et carte des enjeux) induisent une forte variabilité dans l'estimation de l'indicateur DEMA à grande échelle, alors que ce sont d'autres variables, non distribuées spatialement (périodes de retour des crues et courbes d'endommagement), qui expliquent la plus grande part de la variance de l'indicateur DEMA à plus petite échelle. L'approche proposée est applicable à un large éventail de modèles spatialisés, à la condition qu'ils soient peu coûteux en temps de calcul. Elle offre de nouvelles perspectives dans la compréhension des liens entre effets d'échelle, incertitudes et sensibilité, questions qui sont propres à ce type de modèles. Des travaux complémentaires sont à mener pour préciser les effets de biais induit

par la méthode de représentation de l'incertitude sur les variables spatialisées, pour optimiser l'échantillonnage de réalisations aléatoires de ces variables, ou encore pour mieux comprendre et prédire la structure spatiale des cartes de sensibilité.

Références

- [1] M. BIANCHI et C.M. ZHENG : SGeMS : a free and versatile tool for three-dimensional geostatistical applications. *Ground water*, 47:8–12, 2009.
- [2] A. CASTRIGNANO : Accuracy assessment of digital elevation model using stochastic simulation. In *7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, 2006.
- [3] M. CROSETTO et S. TARANTOLA : Uncertainty and sensitivity analysis : tools for GIS-based model implementation. *International Journal of Geographical Information Science*, 15:415–437, 2001.
- [4] M. DELGADO et J. SENDRA : Sensitivity analysis in multicriteria spatial decision-making : a review. *Human and Ecological Risk Assessment*, 10(6):1173–1187, 2004.
- [5] C. V. DEUTSCH et A. G. JOURNEL : *GSLIB : Geostatistical Software Library and User's Guide*. Oxford University Press, New York, 1998.
- [6] K. ERDLLENBRUCH, V. GERMANO, E. GILBERT, F. GRELOT et C. LESCOULIER : Etude socio-économique des inondations sur le bassin versant de l'orb. Rapport technique, Conseil Général de l'Hérault, Programme européen INTERREG III C Sud Europe - INUNDA, 2007.
- [7] K. ERDLLENBRUCH, E. GILBERT, F. GRELOT et C. LESCOULIER : Une analyse coût-bénéfice spatialisée de la protection contre les inondations : Application de la méthode des dommages évités à la basse vallée de l'Orb. *Ingénieries Eau-Agriculture-Territoires*, 53:3–20, 2008.
- [8] S. HESSION, A. M. SHORTRIDGE et N. M. TORBICK : Categorical models for spatial data uncertainty. In *7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, 2006.
- [9] G.B.M. HEUVELINK, S.L.G.E. BURGERS, A. TIKTAK et F.V. DEN BERG : Uncertainty and stochastic sensitivity analysis of the GeoPEARL pesticide leaching model. *Geoderma*, 155(3-4):186–192, 2010.
- [10] B. IOOSS et M. RIBATET : Global sensitivity analysis of computer models with functional inputs. *Reliability Engineering and System Safety*, 94:1194 – 1204, 2009.
- [11] M. LAMBONI : *Analyse de sensibilité pour les modèles dynamiques utilisés en agronomie et environnement*. Thèse de doctorat, Institut des Sciences et Industries du Vivant et de l'Environnement (AgroParisTech), 2009.
- [12] M. LAMBONI, H. MONOD et D. MAKOWSKI : Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models. *Reliability Engineering and System Safety*, 2010. Article in press.
- [13] L. LILBURNE et S. TARANTOLA : Sensitivity analysis of spatial models. *International Journal of Geographical Information Science*, 23(2):151–168, 2009.
- [14] A. MARREL, B. IOOSS, M. JULLIEN, B. LAURENT et E. VOLKOVA : Spatial global sensitivity analysis. Dépôt n°00430171 sur HAL, 2009.
- [15] J. C. REFSGAARD, J. P. van der SLUIJS, A. L. HØJBERG et P. A. VANROLLEGHEM : Uncertainty in the environmental modelling process : A framework and guidance. *Environmental Modelling and Software*, 22:1543–1556, 2007.
- [16] P. RUFFO, L. BAZZANA, A. CONSONNI, A. CORRADI, A. SALTELLI et S. TARANTOLA : Hydrocarbon exploration risk evaluation through uncertainty and sensitivity analysis techniques. *Reliability Engineering and System Safety*, 91(10-11):1155–1162, 2006.
- [17] N. SAINT-GEOURS, J.-S. BAILLY, F. GRELOT et C. LAVERGNE : Is there room to optimise the use of geostatistical simulations for sensitivity analysis of spatially distributed models ? In *Ninth International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, août 2010. Dépôt n°00470529 sur HAL.
- [18] A. SALTELLI, P. ANNONI, I. AZZINI, F. CAMPOLONGO, M. RATTO et S. TARANTOLA : Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2):259 – 270, 2010.

Soumis au Journal de la Société Française de Statistique

File: papierSFds_final.tex, compiled with jsfds, version : 2009/12/09

date: 15 février 2011

- [19] A. SALTELLI, M. RATTO, T. ANDRES, F. CAMPOLONGO, J. CARIBONI, D. GATELLI, M. SAISANA et S. TARANTOLA : *Global Sensitivity Analysis - The Primer*. Wiley, 2008.
- [20] C. SCHEIDT et J. CAERS : Representing spatial uncertainty using distances and kernels. *Mathematical Geosciences*, 41(4):397–419, 2009.
- [21] I. M. SOBOL' : Sensitivity analysis for non-linear mathematical models. *Mathematical Modelling and Computational Experiment*, 1:407–414, 1993.
- [22] E. VOLKOVA, B. IOOSS et F. VAN DORPE : Global sensitivity analysis for a numerical model of radionuclide migration from the RRC Kurchatov Institute radwaste disposal site. *Stochastic Environmental Research and Risk Assessment*, 22(1):17–31, 2008.
- [23] S. P. WECHSLER : Uncertainties associated with digital elevation models for hydrologic applications : A review. *Hydrology and Earth System Sciences*, 11(4):1481–1500, 2007.